

# Autokorrelation von akustischen Sprach- und Musik-Signalen

Keßler, Franz Rudolf

Veröffentlicht in:  
Abhandlungen der Braunschweigischen  
Wissenschaftlichen Gesellschaft Band 51, 2001,  
S.77-92



J. Cramer Verlag, Braunschweig

# **Autokorrelation von akustischen Sprach- und Musik-Signalen**

von **Franz Rudolf Keßler**, Düren\*

(Eingegangen 22.11.2001)

## **Gliederung**

- I Einleitung
- II Resultate der AKA an Beispielen
- III Zeitabhängige Variation der Autokorrelation
- IV Zusammenfassung
- Abbildungen 1 - 16
- Legenden zu den Abbildungen
- Literaturverzeichnis

## **I Einleitung**

In fast allen Teilbereichen der Physik, Informatik und Signal-Meßtechnik stellt die Analyse der Zeitabhängigkeit der Signalfunktion in Bezug auf das inhärente Autokorrelationsverhalten ein wichtiges Verfahren dar, um spezifische „Muster“ und damit Charakteristika in der zeitlichen Abfolge des detektierten Signals zu erkennen.<sup>1</sup> Unter der Zeitfunktion eines akustischen Signals ist je nach Fragestellung z.B. die Zeitfunktion des Schalldrucks, die elektrische Ausgangsspannung eines HiFi-Verstärkers oder die momentane abgegebene Leistung einer Schallquelle zu verstehen. Das zur Analyse gelangende Signal kann dabei auch aus der Überlagerung mehrerer Signalanteile und/oder von Rauschbeiträgen bestehen. Bei der Interpretation der Ergebnisse der Autokorrelationsanalyse (AKA) müssen selbstverständlich alle beteiligten Glieder der akustischen Signalausbreitung von der (Schall-)Quelle bis hin zur (Schall-)Detektion unter Einbezug der u. U. modifizierend wirkenden Einflüsse längs des Übertragungsweges berücksichtigt werden.

Unter „Autokorrelation“ einer Zeitfunktion versteht man den zeitlich systematischen, also gewissermaßen „inneren“ Zusammenhang zwischen zwei Werten der Größe des Signals, die in einem zeitlichen Abstand „ $\tau$ “ ermittelt werden. Vorausgesetzt ist dabei, daß der Mittelwert der Signalgröße, genommen über einen hinreichend großen Zeitraum, Null,

---

\* Univ.Prof. em. Dr. Franz Rudolf Keßler, Inst. für Halbleiterphysik und Optik der Techn. Universität Carolo Wilhelmina zu Braunschweig

Privatanschrift: Am Krausberg 12 · D-52351 Düren · e-mail: FranzRudolf.Kessler@t-online

<sup>1</sup> Sogar das Autokorrelationsverhalten der Zeitfunktion von Aktienkursen ist bereits untersucht worden [C2].

zumindest aber hinreichend stationär ist. Man betrachtet zur Beurteilung der inhärenten Korrelation die Summe aller möglichen Produkte  $P_i(\tau)$  von zwei Signalwerten, die einen Abstand  $\tau$  besitzen. Normiert man die Summe der Produkte auf die Zeitspanne, über der das Signal in dieser Weise ausgewertet wird (das ist die sogen. „Integrationszeit“), so ist  $(\sum P_i(\tau))_{\text{norm}} = f(\tau)$  die gesuchte Autokorrelationsfunktion (AKF) in Abhängigkeit von der „Verzögerungszeit“  $\tau$ . Die Grundlagen der Autokorrelationsanalyse (AKA) sind sowohl in Bezug auf die theoretischen Prinzipien als auch in Bezug auf die meßtechnische Realisierung, sei diese apparativ oder Computer-gestützt, an vielen Stellen der zugänglichen Literatur detailliert dargestellt, so daß hier auf die dortigen Ausführungen verwiesen werden kann (s. z.B. [A6]).

Erkenntnisse über bzw. Hinweise auf das Autokorrelationsverhalten speziell von Sprache und Musik sind in der Literatur vergleichsweise kaum zu finden, obwohl sich in dem Zeitablauf von Sprache und Musik fraglos recht interessante Charakteristika finden, die den Autokorrelationstypus beeinflussen und sich überdies ganz verschiedenen wissenschaftlichen und praktischen Disziplinen der Sprach- und Musiktheorie zuordnen lassen. Verbindungen bestehen z. B. im Bereich der Sprachanalyse zur Linguistik, Rhetorik, logopädischen Therapeutik und zur Charakterisierung stilistisch spezifischer Satzkonstruktionen. Im Bereich der Musik sind mit der AKA z.B. Fragen der Melodieführung, der Rhythmusgestaltung, der Instrumentalisierung, der Komponisten-Interpretation, der Zuordnung zu bestimmten Stilarten oder Genre-Attitüden usw. zu klären [C3].

Eine mögliche erste Fragestellung ist z. B. auch die nach dem „Empfindungsunterschied“ von Sprache und Musik selbst. Jedermann kann nach wenigen Zehntelsekunden des Zuhörens sagen, ob da Musik oder Sprache aus dem Lautsprecher klingt. Neben den im Signal enthaltenen Frequenzen des akustischen Spektralbereiches spielt dabei auch das Autokorrelationsverhalten in der zeitlichen Ton- und/oder Intensitätsabfolge, gemeint ist in letzterem Fall die zeitlich variierende Betonung durch Lautstärke- und/oder Tempi-Wechsel, eine Rolle. Z. B. lehrt ja auch die Erfahrung, daß man eine bekannte oder unbekannte Melodie erst dann richtig „erfassen“ kann, d.h. wiedererkennen bzw. als „sinnvoll“, „schön“ oder in sich harmonisch einschätzen kann, wenn die entsprechende Passage der Tonempfindung in ihrer Abfolge im Gehirn eine bestimmte Zeit lang, d. h. etwa 5 bis 10 Sekunden lang, gespeichert und dabei als Ganzes, d.h. als eine typische „Mikro-Einheit“ des Musikstückes „bewertet“ worden ist. Das Gehirn vollzieht dabei, natürlich unbewußt, fraglos auch eine Art „Autokorrelationsanalyse“.

Man könnte sich, um den Hinweis einer kleinen praktischen Anwendung der Sprach- und Musikanalyse ins Spiel zu bringen, z. B. auf der Basis einer online-AKA, d.h. mit dem Resultat einer permanenten und hinreichend schnellen Diskriminierung zwischen Sprache und Musik mittels AKA, ein elektronisches Hilfsgerät vorstellen, das für den Lautsprecher oder den Recorder je nach Wunsch entweder den Sprach- oder den Musik-Teil einer laufenden Rundfunk- oder Fernsehsendung lauter oder leiser stellt bzw. stumm schaltet.

Wegen der Beteiligung des menschlichen Hörorgans und des Gehirns, die beide jeweils eine zusätzliche individuelle „Filterung“ des Signals bewirken, an der Empfindung bzw. an der intellektuellen Bewertung von Sprache und Musik spricht man auch von Psychoakustik bzw. im englischen Sprachraum (s. z.B. [B5]) von „psychoacoustics“ oder auch „head

acoustics“ als dem relevanten Spezialgebiet der Akustik, dem also auch die AKA der Sprach- und Musiksiknale zuzuordnen ist. – Auf die völligen Analogien der diesbezüglichen Signalverarbeitung bzw. -analyse in Bezug auf den Signalaustausch im Tierreich sei hier nur kurz hingewiesen.

Die Computer-gestützte AKA an einer Sprach- oder Musik-Passage kann zunächst „nachträglich“, d. h. an einer festgelegten Anzahl von Meßwerten stattfinden, die zuvor in eine numerische Matrix eingestellt wurden. Die maximale Integrationszeit ist damit festgelegt. Die AKA kann aber auch kontinuierlich, d.h. „online“ erfolgen, indem laufend weitere Meßwerte aufgenommen werden. In letzterem Fall liegt dann eine sich zeitlich kontinuierlich ausdehnende Integrationszeit bzw. eine stärker „gestützte“ Mittelwertbildung für die zu bestimmende Autokorrelationsfunktion (AKF) vor.

Die AKF kann drittens auch aus einer jeweils begrenzten Zahl von Meßwerten gewonnen werden, die zwar eine zweckmäßig gewählte Integrationszeit erlauben, dann aber zyklisch „regeneriert“ werden, so daß zeitliche Veränderungen in der AKF wenigstens schrittweise verfolgt werden können. Diese Form einer kontinuierlich repetierten AKA empfiehlt sich insbesondere in den Fällen eines zeitlich nicht hinreichend stationären Signalmittelwertes und/oder in Fällen des Vorliegens zeitabhängiger Veränderungen in den Korrelationseigenschaften des Signals selbst.

Die im folgenden mitgeteilten Ergebnisse an Sprach- und Musikbeispielen beziehen sich auf Computer-gestützte AKA-Messungen der elektrischen Spannungssignale am Ausgang eines linear arbeitenden Meßgleichrichters, an dessen Eingang die Ausgangsspannung eines HiFi-Verstärkers liegt. Mittels eines in seiner Selektivität spezifisch einstellbaren Equalizers [Fa. Technics, Typ SH9010] wird u. U., wie dann gesondert angegeben, eine Frequenzfilterung vorgenommen. Z. B. kann ein schmalbandiges Durchlaßfilter mit der Schwerpunktfrequenz  $f_0$  zwischen 30 Hz und 25 kHz mit einer 3-dB-Bandbreite  $\Delta f$  bis herab zu  $\Delta f = (f_0 / 7)$  Hz zwischengeschaltet werden.

Die untersuchten und hier nur in einer charakteristischen Auswahl dargestellten Sprach- und Musik-Muster wurden selbst erstellt oder laufenden Rundfunksendungen bzw. abgespielten kommerziellen Tonträgern entnommen. – Die in diesem Sinne durchgeführten Autokorrelationsanalysen wurden mittels selbst entwickelter Basic-Programme auf einem eigenen PC mit Pentium-Prozessor ausgeführt.

Das beigelegte Literaturverzeichnis enthält allgemeine Hinweise auf die AKA-Methode und zitiert spezielle Arbeiten zur AKA akustischer Signale. Aufgenommen wurden auch einige grundlegende Werke zur akustischen Analyse bzw. theoretischen Behandlung von Sprach- und Musiksignalen.

## II Resultate der AKA an Beispielen

Der hier verwendete numerische Algorithmus zur Bestimmung der Autokorrelationsfunktion (AKF) mittels eines Personal Computers ist in einer vorangegangenen Arbeit des Autors ausführlich dargestellt (s. [A6]). Das Meßmodul zur Wandlung des analogen Spannungssignals in digitale Werte zur Programmverarbeitung erlaubt konstruktionsbedingt nur einen kleinsten Zeitschritt von ca.  $\delta t = 0,15$  [s]. Dementsprechend ist die ausgegebene

AKF die AKF der Einhüllenden Amplitudenfunktion  $A_0(t)$  des niederfrequenten akustischen Signals, denn der Autokorrelator kann den quasi-sinusförmigen Spannungsamplituden im akustischen Frequenzbereich von etwa 40 Hz bis 25 kHz nicht folgen, registriert demnach nur den zeitabhängigen Effektivwert der Spannung.

Die erhaltenen Ergebnisse der AKA an einigen Sprach- und Musik-Beispielen werden im folgenden in der Form von „Thesen“ aufgeführt. Jeweils erfolgt dazu der Hinweis auf die entsprechenden Abbildungen, die die ermittelte AKF zeigen und als „Beleg“ gelten sollen. Die Abbildungen sind aus Gründen der Platzersparnis im Format möglichst klein gehalten, gerade so, daß man den prinzipiellen Verlauf der AKF erkennen kann. Die Entnahme von Zahlenwerten aus den Abbildungen ist also nicht vorgesehen. Die jeweils ausführlichen Legenden zu den Abbildungen sind als Textseiten zusammengefaßt. Dort können die zu einer Abbildung gehörigen spezifischen Parameter entnommen werden. Durch diese Separierung von Abbildung und ausführlicher Legende rücken die Abbildungen mit Gewinn für deren Vergleich enger zusammen.

Im Einzelnen kann festgestellt werden:

**These 1:** Die AKF von Sprache und Musik verläuft dem Grunde nach ähnlich, d.h. nach einem quasi exponentiellen Abfall folgt ein relativ flach verlaufender Teil, in dem nach einem Minimum i. allg. entweder ein schwacher und breiter Wiederanstieg oder eine quasi periodische Modulation zu erkennen ist. Als „Beleg“ für diese Aussage enthalten die Abb. 1 bis 10 eine Reihe unterschiedlicher AKF-Verläufe von stilistisch ganz verschiedenen Sprach- und Musikbeispielen, die in den folgen Thesen noch differenzierter verglichen werden.

**These 2:** Der exponentielle Relaxationsparameter  $\tau_{1/e}$  im Anfangsteil der AKF ist bei Sprache und Musik signifikant verschieden und beträgt bei Sprache etwa  $\tau_{1/e} = 1$  [s] und bei Musik i. allg. mehr als  $\tau_{1/e} = 2$  [s]. Der Wert bei Sprache korrespondiert mit der häufigsten zeitlichen Wortlänge in Verbindung mit der zeitlichen Halbwertsbreite der statistischen Wortlängenverteilung bei gegebenem Sprecher, Textautor und Sprachtypus. Der Abfallparameter bei der Musik ist durch zeitliche Länge der rhythmisch dominierenden Taktgruppe bzw. der Melodie des Musikstückes bestimmt und damit ebenso abhängig vom Komponisten, wie vom Dirigenten der Musik-Reproduktion. In der Regel ist die zeitliche Länge der prominenten Struktur im statistischen Mittel bei Musik offenbar mehr als doppelt so groß, als bei Sprache.

Der signifikant verschiedene exponentielle Anfangs-Relaxationsparameter, der also für kleine Verzögerungszeiten  $t$  in der AKF bestimmend ist, kann insbesondere durch Vergleich der in den Abb. 1 und 2 für Musik einerseits und in den Abb. 3 und 4 für Sprache andererseits gegebenen Beispiele bestätigt gefunden werden. – Zur Wortlängenstatistik siehe die Legende zu Abb. 5.

**These 3:** Der in These 1 konstatierte Wiederanstieg in der AKF hängt bei der Sprache mit der Satzlänge, bzw. deren statistischer Verteilung im konkreten Fall ab. Der gesprochene Satz besitzt nämlich eine signifikante Lautstärke-Verlaufsstruktur. Durch die quasi-periodische Wiederholung, wenn auch mit statistischer Schwankung der Länge dieser Struktur entsprechend der Schwankung der Satzlänge, tritt, wie auch theoretisch zu verste-

hen, ein praktisch nur einmaliger Wiederanstieg in der AKF auf (s. dazu Abb. 4 und 5). – Zur konkreten Wort- und Satzlängen-Verteilung in dem herangezogenen Text für die in Abb. 5 dargestellte AKA siehe die Ausführungen in der entsprechenden Abbildungslegende.

Bei der Musik (s. Abb. 2) tritt an die Stelle der Satzlänge eine Betonungsgruppierung, d.h. eine entsprechende Lautstärkevariation, die, vom Komponisten bzw. Dirigenten vorgegeben, das Verständnis der Musik erleichtern und/oder ein dramatisches Element in deren Ablauf bringen soll. Auch hier schwankt die Länge, sie ist jedoch nicht unkorreliert statistisch verteilt. Eine derartige Betonungsstruktur ist bekanntlich bei klassischer Musik noch ausgeprägter vorhanden, als bei Tanzmusik.

**These 4:** Treten in der Sprache oder in der Musik durch einen ausgeprägten Wort- bzw. Takt-Rhythmus signifikant periodische Strukturen auf, so verläuft die AKF auch mit einem signifikant oszillatorischen Anteil gleicher Periodendauer. Als Beleg dafür wurde in einem Sprechversuch das Wort „Panamalkanal“ im gleichen Tonfall periodisch wiederholt. Die zugehörige AKF ist in Abb. 8 wiedergegeben. – Als diesbezügliches Musik-Beispiel wurde der „Boléro“ von M. Ravel gewählt. Die zugehörige AKF ist in Abb. 7 wiedergegeben. – Auch das Musikbeispiel des Klavierstückes „Danses gothique“ von E. Satie zeigt infolge des rhythmischen Klavierspiels entsprechende Oszillationen: Abb. 9.

**These 5:** Die gefundene AKF ist bei Sprache nicht signifikant verschieden, wenn einmal eine Frau und einmal ein Mann vorliest. Dies gilt natürlich nur für den gleichen oder wenigstens „in der Art“ gleichen analysierten Sprechtext. Vorausgesetzt ist dabei ferner, daß der Vortragsstil als solcher ähnlich ist. D. h. es sollen keine „dramatischen“ Unterschiede in der Betonung und der Pausengestaltung bestehen. Man vergleiche dazu die Abb. 3 und 4. In beiden Fällen handelt es sich um verschiedene Nachrichtentexte des Deutschlandfunks (DLF).

**These 6:** Klassische Musik und Tanz- bzw. Schlagermusik zeigen in der AKF gewisse Unterschiede, wobei jedoch wiederum von Komponist zu Komponist signifikante Differenzierungen auftreten. Die quantitative Zuordnung dieser Unterschiede zu bestimmten Stil- und Spielarten würde jedoch eine viel breiter angelegte „Feldstudie“ der AKA an einem systematisch variierten Vorlagenmaterial voraussetzen. – Um jedoch hier den interessierten Hörer auf existierende Unterschiede hinzuweisen, sei der Vergleich der folgenden Beispiele empfohlen: Abb. 1, 2, 7 und 9.

**These 7:** Eine spektrale Filterung des akustischen Signals verändert die AKF in signifikanter Weise. Dies gilt tendenziell sowohl für Sprache, als auch für Musik, wobei die Veränderungen wesentlich in den Bereichen größerer Verzögerungszeit, d.h. deutlich oberhalb der Anfangsrelaxationszeit auftreten:

- a) Sprache: Bei z.B. Hochpassfilterung werden quas-periodische Anteile in der AKF aufgrund der Satz-Betonungsmodulation in der Lautstärke bzw. die niederfrequenten Stimmanteile reduziert zugunsten des erweiterten Hervortretens des exponentiellen Anfangsabfalls der AKF, wie er als Folge der Wortlängen- und Wortklangstatistik auftritt. – Diesbezüglich sei auf einen Vergleich zwischen den Abb. 5 und 6 verwiesen. Das für beide AKA-Fälle verwendete Sprech-Beispiel ist primär das gleiche.

b) Musik: Auch bei der Musik wird offenbar als Folge der Hochpaßfilterung der durch den inhärenten Rhythmus bedingte Wiederanstieg der AKF nach dem anfänglich signifikant ausgeprägten exponentiellen Abfall reduziert bzw. eliminiert, so daß der exponentielle Abfall praktisch monoton bis auf den stationären Endwert erfolgt.

Dieses Verhalten ist in Abb. 10 gut quantitativ zu verfolgen, denn bei dieser Darstellung der AKF waren die Meßwerte vorher durch Subtraktion des Mittelwertes mittelwertfrei gemacht worden. Bei mittelwertfreien Meßwerten ist der asymptotisch stationäre Endwert stets gleich Null. – Ferner wurde die AKF auf ihren Maximalwert normiert (Näheres dazu s. These 8). Damit liegen hier sämtliche Werte der AKF zwischen Eins und Null.

**These 8:** Zur Bestimmung der Charakteristika von Sprach- bzw. Musikmustern erscheint es zweckmäßig, im konkreten Fall neben der diesbezüglichen AKF selbst noch zwei weitere Funktionen ins Auge fassen:

1.) Die zeitliche Abfolge der Meßwertgröße in der Abhängigkeit von der Zeit bzw. von der Zahl der vergangenen Zeitschritte  $\delta t$ .

Die zeitliche Abfolge der Meßwerte bestimmt zwar eindeutig die AKA, das gilt jedoch nicht umgekehrt. Z. B. kann eine exponentielle Funktion der AKF durch ganz unterschiedliche funktionale Meßwertabfolgen im Zeitbereich erzeugt sein. Die Kenntnis der zeitlichen Abfolge der Meßwertgrößen enthält also noch eine eigenständige Information über die Charakteristika des analysierten Sprach- bzw. Musikmusters. Außerdem ermöglicht bereits die graphische Auftragung der Meßwerte als Funktion der Zeit eine zusätzliche visuelle Einschätzung:

- a) ob der Mittelwert in relevanten Zeitabschnitten  $\Delta t$  hinreichend konstant ist,
- b) ob sich eine quasi-periodische Modulation erkennen läßt und
- c) ob sich eine größere bzw. signifikante Zeitstruktur in Wiederholung zeigt.

Im Fall a) macht es i. allg. der unzulässig schwankende Mittelwert erforderlich, die gespeicherten Meßdaten in der Matrix einer rechnerischen Normierung zu unterziehen, so daß relativ sehr langsame Mittelwertschwankungen im Sinne eines Fadingausgleichs eliminiert werden. Andererseits kann die Feststellung einer unzulässigen Mittelwertschwankung auch Veranlassung sein, bereits eine apparativ-elektronische Fading-Regelung mit entsprechender Zeitkonstanten grundsätzlich der Signaldetektion vorzuschalten.

Auf die Möglichkeit, bei *konstantem* Mittelwert die Meßwerte durch Subtraktion des Mittelwertes „mittelwertfrei“ zu machen und so eine leichter vergleichbare Darstellung der Ergebnisse zu gewinnen, wurde oben bereits hingewiesen. Damit reduziert sich nämlich der asymptotische Grenzwert der AKF für  $\tau \rightarrow \infty$  auf Null. Ferner kann in diesem Fall die AKF neben der stets stattfindenden Normierung auf die Integrationszeit 1 auch noch auf den Wert bei  $\tau = 0$ , das ist die Varianz der mittelwertfreien Meßwerte, normiert werden, so daß gilt:  $AKF_{\text{norm}}(\tau = 0) = 1$ . Wie theoretisch bewiesen ist, ist der Wert der Varianz der mittelwertfreien Meßwerte stets der Maximalwert der AKF. – In dieser Weise war in der AKA zu Abb. {10} in These 7 verfahren worden.

2.) Die Häufigkeitsverteilung der Meßwertgröße bezogen auf die Gesamtheit der registrierten Meßwerte.



Die Kenntnis der Häufigkeitsverteilung der Meßwertgrößen vermittelt im konkreten Fall ebenfalls eine erweiterte Einsicht in die „innere Struktur“ des detektierten akustischen Signals. Diese zusätzliche Kenntnis ist wiederum insofern eigenständig, als die Häufigkeitsverteilung der Meßwerte zwar ein Charakteristikum der Signalfunktion ist, jedoch der spezifische funktionale Verlauf der AKF nicht von ihr abhängt. – Dazu sei bemerkt, daß bei einer beliebigen Häufigkeitsverteilung stets dann die AKF, der Wert bei  $\tau = 0$  ausgenommen, bei beliebigen Zeitverzögerungen konstant gleich Null ist, wenn diese Meßwerte zeitlich völlig unkorreliert nacheinander auftreten. Ein statistisch unkorreliertes Rauschsignal führt, um ein Beispiel zu nennen, zu eben einer solchen AKF mit dem Wert Null für  $\tau > 0$ . Dabei ist es ohne Einfluß, ob das Rauschen „weiß“ ist oder eine bestimmte *spektrale* Amplitudenverteilung besitzt.

Um die AKA in diesem Sinne zu erläutern bzw. zu ergänzen, wurde als Beispiel zu der in Abb. 3 gegebenen AKF in Abb. 11 die zugehörige Zeitfunktion der Meßwerte und in Abb. 12 deren Häufigkeitsverteilung dargestellt. – Auf die in Abb. 11 erkennbaren Sprechpausen sei hier nur am Rande hingewiesen.

Zwei weitere Serien von „effektiven Meßwerten“ in ihrer zeitlichen Abfolge sind in Abb. 13 und Abb. 15 dargestellt, sie werden im folgenden Abschnitt III im Zusammenhang mit der „kontinuierlichen AKA“ näher erläutert.

### III Zeitabhängige Variation der Autokorrelation

Ist die Zeitfunktion des Signals entweder in ihrem Mittelwert und/oder in ihrer inhärenten Autokorrelationseigenschaft zeitlich nicht stationär, so ist die Autokorrelationsfunktion (AKF) in ihrem Verlauf ebenfalls zeitabhängig: Das Relaxationsverhalten variiert dabei i. allg. in seinen Parametern, aber auch der funktionale Typus der AKF kann sich zeitlich ändern. Da der asymptotische Grenzwert der AKF dem Quadrat des Mittelwertes der Signalfunktion entspricht, zieht die zeitliche Änderung des Mittelwertes eine Änderung des Bezugs-Nullpunktes für das Abklingen der AKF nach sich. Liegen solche Fälle vor, bzw. ist eine entsprechende zeitliche Variation zu erwarten, so muß eine diesbezügliche Prüfung stattfinden, möglichst in Verbindung mit einer quantitativen Analyse. – Hier bietet sich die „kontinuierliche Autokorrelationsanalyse“ (KONTAKA) an. Zur Überwachung genügt u.U. bereits die laufende Beobachtung, d.h. die Anzeige, der Differenz  $\Delta[\text{AKF}, \Delta\tau]$  zweier Autokorrelationskoeffizienten für eine geeignet gewählte Zeitdifferenz  $\Delta\tau$  in der Verzögerungszeit.

Die KONTAKA arbeitet dabei mit einer festen Anzahl  $Z$  von Meßwerten in einem „Meßdatenregister“. Der Wert  $\Delta[\text{AKF}, \Delta\tau]$ , mit  $\Delta\tau = (\tau_2 - \tau_1)$ , wird gebildet aus den beiden relevanten AKF-Werten mit  $\tau_1$  und  $\tau_2$  (mit  $\tau_2 > \tau_1$ ). Die größten Differenzen erhält man i.allg., z. B. bei exponentieller AKF-Relaxation, mit dem Wert  $\tau_1 = 0$ . Jeder AKF-Wert entsteht aus der Aufsummation aller möglichen Meßwert-Produkte für die feste Verzögerungszeit  $\tau_1$  bzw.  $\tau_2$ . Ist  $\tau_2 = M \cdot \delta t$ , so wird also  $\Delta[\text{AKF}, \Delta\tau]$  als normierter Wert einer Produktaddition über  $(Z-M)$  Meßwerte, bzw. über eine „Integrationszeit“ von  $(Z-M) \cdot \delta t$ , gebildet. Meßtechnisch gesehen, d. h. zur „Verbesserung“ des Resultates der Integration, sollte  $Z$  deswegen hinreichend viel größer als  $M$  gewählt werden.



Ist auf diese Weise  $\Delta[\text{AKF}, \Delta\tau, t_1]$  zur Zeit  $t_1$  gewonnen, so wird der erste Meßwert im Register gelöscht, dann werden die Meßwerte geschlossen um einen Platz nach vorne verschoben und schließlich wird der bisher letzte Meßwert durch den nächsten neuen Meßwert, zur Zeit  $(t_1 + \delta t)$  aufgenommen, ersetzt. – Dieser Zyklus setzt sich ununterbrochen fort, so daß auf diese Weise die Funktion

$$\Delta[\text{AKF}, \Delta\tau, t_n] = f(t_n) = f(t_1 + n \cdot \delta t) \text{ mit } n = 0, 1, 2, 3, \dots \quad \text{Gl. (III,1)}$$

gewonnen wird.

Wegen des (mutmaßlich oder konkret gegebenen) nicht stationären Mittelwertes der Signalfunktion, muß man vor der Produktbildung zur AKF alle Meßwerte im Register auf einen festen Mittelwert normieren. Nur dann bezieht sich der jeweilig erhaltene Wert nach Gl. (III,1) mit jedem Zeitschritt auf den gleichen stationären Endwert der AKF, d. h. auf  $\text{AKF}(\tau \rightarrow \infty)$  und damit auf das Quadrat dieses „stabilisierten“ Mittelwertes.

Um die Differenz  $\Delta[\text{AKF}, \Delta\tau, t_n]$  noch signifikanter darzustellen, ist es zweckmäßig, sie ausschließlich auf den Relaxationsteil der AKF zu beziehen. Man bildet deshalb besser die relative Differenz des Abstandes zum asymptotischen Grenzwert, d.h. man wählt zur Beurteilung des Zeitverhaltens der AKF die Darstellung der Größe:

$$Y(t_n) = \Delta[\text{AKF}, \Delta\tau, t_n] / [\text{AKF}(\tau_1) - \text{AKF}(\tau \rightarrow \infty)] \quad , \quad \text{Gl. (III,2)}$$

deren Wertebereich zwischen Null und Eins bzw. 0 und 100 % liegt.

Aufgrund der im vorangegangenen Abschnitt präsentierten Ergebnisse der AKA von Sprach- und Musikmustern ist zu erwarten, daß für relativ kleine Verzögerungszeitdifferenzen  $\Delta\tau$ , aufgenommen mit  $\tau_1 = 0$ , der zeitabhängige Wert  $Y(t_n)$  [definiert in Gl. (III,2)] bei Sprache generell signifikant größer ist als bei Musik, denn die Zeitkonstante der anfänglichen exponentiellen Relaxation der AKF ist bei Sprache signifikant kleiner als bei Musik.

Die Abb. 14 und 16 zeigen das Meßergebnis, d.h. die zeitabhängige Variation der in Gl. (III,2) definierten Größe  $Y(t_n)$ , für je ein Musik- bzw. Sprachbeispiel. Trotz relativ großer Schwankungen, die die zeitabhängige „Dramatisierung“ bzw. Phrasierung der Musik bzw. der Sprache zum Ausdruck bringen, findet sich die Erwartung deutlich bestätigt: Für den Wert  $\Delta\tau = 1,26$  [s] ergibt sich bei einer Integrationszeit von 413 [s] bei ausgewerteten 2000 effektiven Meßwerten (entsprechen einer jeweiligen Aufnahmezeit von 360 [s] = 6') bei der Musik ein Mittelwert der Größe  $Y(t_n)$  [nach Gl. (III,2)] von 5,3 % und bei Sprache von 10,8 %, also ein um den Faktor 2,04 größerer Wert.

Die Abb. 13 und 15 stellen die Größe des „ersten korrigierten Meßwertes“ im Register für die jeweilige Bestimmung der AKF-Differenz in der zeitlichen Abfolge der Gesamtheit der im Zeitablauf ausgewerteten Meßwerte dar. Man erkennt in beiden Fällen die wirksame längerfristige Mittelwertstabilisierung auf den Wert 10. – Die auf dieser Gesamtheit der Meßwerte basierende Zeitabhängigkeit der AKF-Differenz ist die in den Abb. 14 und 16 dargestellte.

In Abb. 13 ist in dem „Muster“ der Meßwertabfolge deutlich festzustellen, daß das Musikbeispiel aus einem anfänglichen Teilstück von ca. 900 Meßwerten besteht, das sich danach wiederholt. Das entspricht der in diesem Fall absichtlich wiederholt aneinandergefügt

Tonbandaufnahme des Schlagers, um dem Musikbeispiel die notwendige Länge zu geben, ohne den inhärenten Korrelations-Charakter zu verändern. – Die effektiven Meßwerte in Abb. 15 dagegen entsprechen einem fortlaufend vorgelesenen Text ausreichender Länge.

Blickt man auf die in der Einleitung genannte beispielhafte Anwendung einer gesuchten Unterscheidung zwischen Musik und Sprache mittels der kontinuierlichen AKA, so kann diese aufgrund des obigen Resultates als tatsächlich realisierbar eingeschätzt werden. Allerdings muß dabei bedacht werden, daß selbst bei sprungartigem Wechsel zwischen Sprache und Musik das verwertbare Unterscheidungssignal nur nach einer Verzögerungszeit auftritt, die in der Größenordnung der Integrationszeit liegt. Dies ist darin begründet, daß das Register der Meßwerte erst nach dieser Zeitspanne mit den Meßwerten des neuen Signaltypus vollständig gefüllt ist. Erst ab diesem Zeitpunkt entspricht also der ausgegebene  $Y(t_n)$ -Wert [ nach Gl. (III,2) ] dem anliegenden Signaltypus. Eine Verkürzung der Integrationszeit würde andererseits grundsätzlich stärkere Schwankungen der oben für Musik und auch Sprache erwähnten und „in sich“ begründeten Schwankungen des  $Y(t_n)$ -Wertes zur Folge haben.

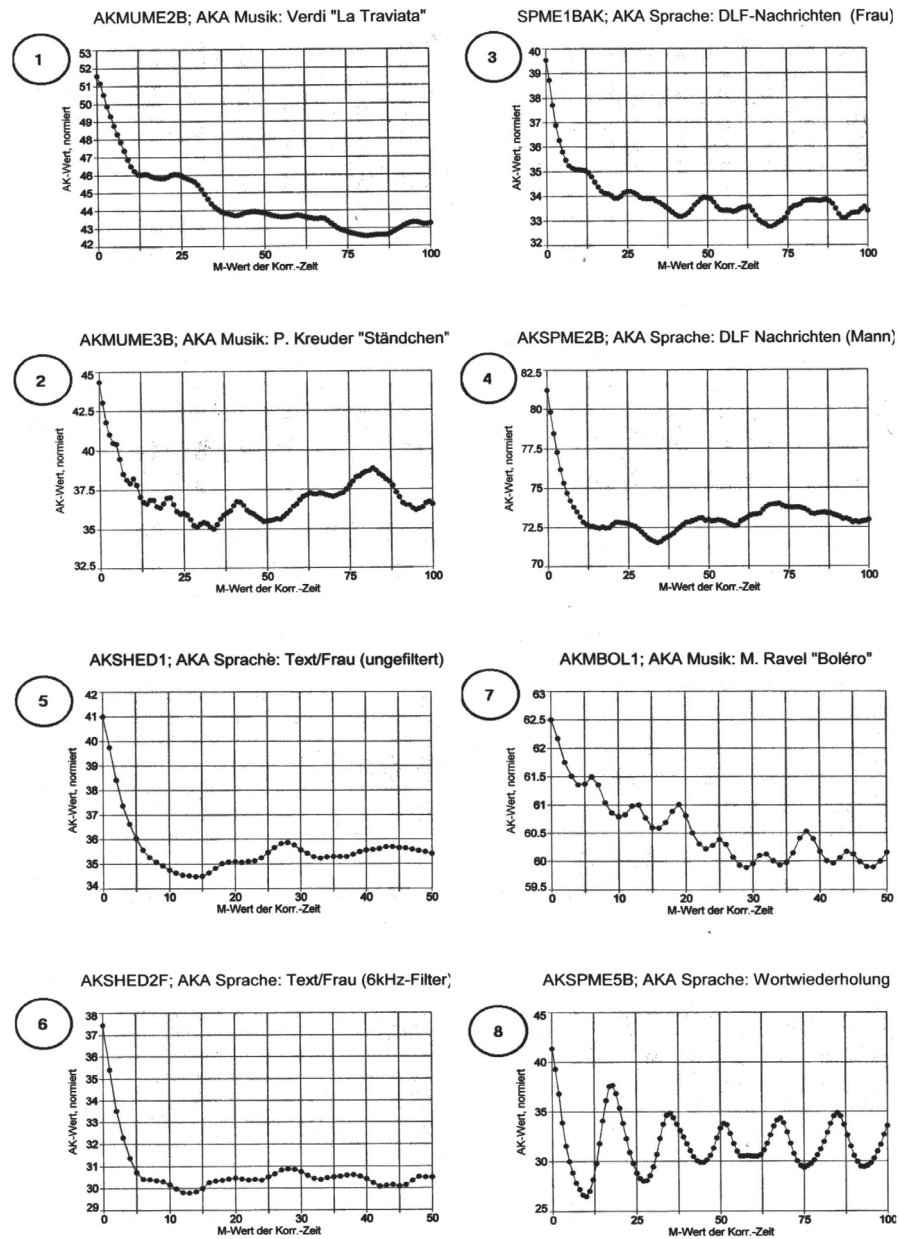
#### IV Zusammenfassung

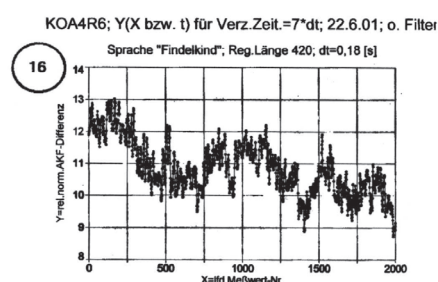
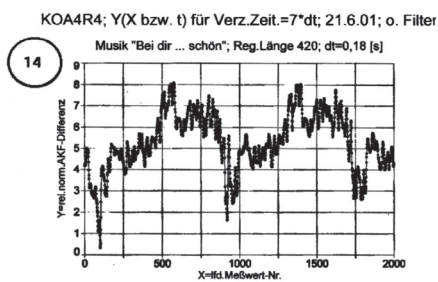
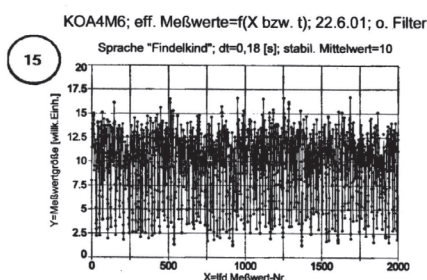
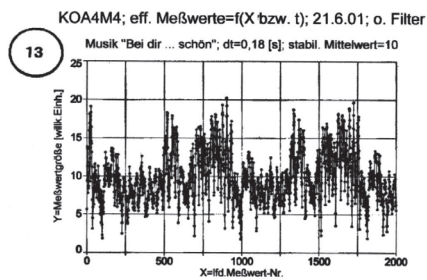
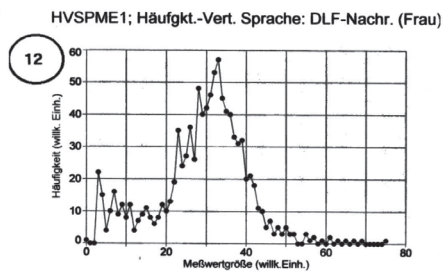
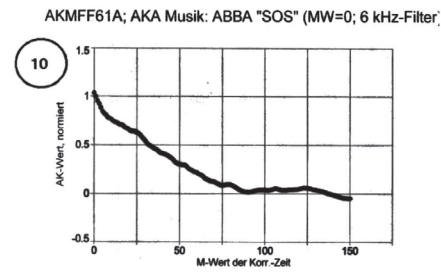
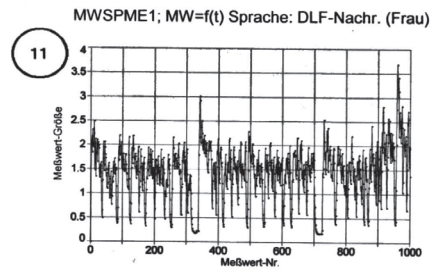
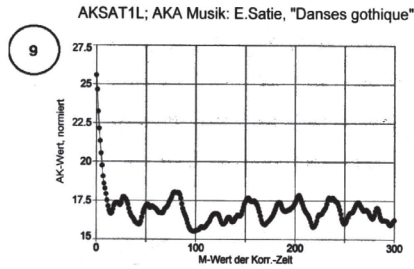
Die vorstehenden Ausführungen haben das Ziel, die Aufmerksamkeit der Leser einmal speziell auf die Autokorrelationsanalyse (AKA) von Sprache und Musik und damit auf ein interessantes Teilgebiet der Signal- bzw. Kommunikationstheorie zu lenken. Im Rahmen der modernen, in vielen Fällen auf komplizierten elektronisch unterstützten Wegen erfolgenden Signalerzeugung, -übertragung und -speicherung sollten die vorgetragenen Aspekte nach Meinung des Autors auf ein breiteres und dabei auch interdisziplinäres Interesse stoßen. Z. B. könnten sich hilfreiche Erkenntnisse aus der AKA der Sprache bzw. Musik für die logopädische Therapeutik einerseits und die Musik-Interpretation andererseits ergeben.

In der vorliegenden Arbeit wurden einige typische Sprach- und Musik-Beispiele der AKA unterzogen, das Ergebnis graphisch dargestellt und kurz diskutiert. Dabei wurde auch auf die ergänzenden Informationen hingewiesen, die aus der zeitlichen Abfolge der Meßwerte und aus der Häufigkeitsverteilung der Größe der Meßwerte gewonnen werden können. – Die wesentlichen Ergebnisse sind in „Thesen“ zusammengefaßt.

In diesem Zusammenhang wird auch die „kontinuierliche AKA“ (KONTAKA) in ihrer Bedeutung für die Erfassung der Zeitabhängigkeit der in Musik- und Sprachdarbietungen u. U. laufend wechselnden Korrelationseigenschaften vorgestellt und anhand eines Anwendungsbeispiels erläutert. Unverzichtbarer Bestandteil der KONTAKA ist die hier entwickelte „Mittelwertstabilisierung“ bei existierender Phrasierung bzw. Lautstärke-schwankung von real dargebotenen Sprach- und Musiksignalen.

Diese hier berichteten Ergebnisse und hinzugefügten kurzen Erläuterungen können nicht vollständig alle Teilaspekte berücksichtigen, zumal die akustisch charakteristischen Parameter einen außerordentlich großen Variationsbereich umfassen. Umso mehr mögen die vorliegenden Ausführungen anregen, entsprechende Untersuchungen im Rahmen systematischer „Feldversuche“ bzw. gezielter Variationen und Erweiterungen der Sprach- und Musikdarbietung fortzuführen und dabei auch durch begleitende theoretische Modellrechnungen auf eine erweiterte quantitative Basis zu stellen.





## Legenden zu den Abbildungen

### Abb. 1: AKMUME2B; AKA Musik: Verdi „La Traviata“

Aus der Oper „La Traviata“ von G. Verdi der Ausschnitt: 1. Akt, Preludium, Teilstück Barcarole“ (4/4 Takt) CD-Produktion „Classical Collection, Vol. 2, Nürnberger Symphoniker, Leitung Alexander von Pitamic © Mediaphon-Madacy, 1999; MJ-22.632 AKA mit 1000 Meßwerten in 145 [s]; Zeitschritt  $\delta t = \delta \tau = 0,145$  [s]

### Abb. 2: AKMUME3B; AKA Musik: P. Kreuder „Ständchen“

Peter Kreuder spielt am Klavier mit seinen Solisten das „Ständchen“ von Heykens ( 2:57 Gesamtdauer) Ausschnitt aus der CD-Produktion „In der Bar nebenan 1935-1944“ Bob CD 17 „Bob's Music“, Hamburg, 1999 [Historische Originalaufnahmen überspielt von Schellackplatten] AKA mit 1000 Meßwerten in 145 [s]; Zeitschritt  $\delta t = \delta \tau = 0,145$  [s]

### Abb. 3: SPME1BAK; AKA Sprache: DLF-Nachrichten (Frau)

Aufnahme der Nachrichten des Deutschlandfunks am 22.2.01; 14:00 Uhr mit Sprecherin; ein beliebiger Ausschnitt; AKA mit 1000 Meßwerten in 145 [s]; Zeitschritt  $\delta t = \delta \tau = 0,145$  [s]

### Abb. 4: AKSPME2B; AKA Sprache: DLF-Nachrichten (Mann)

Aufnahme der Nachrichten des Deutschlandfunks am 25.2.01; 8:00 Uhr mit Sprecher; ein beliebiger Ausschnitt; AKA mit 1000 Meßwerten in 145 [s]; Zeitschritt  $\delta t = \delta \tau = 0,145$  [s]

### Abb. 5: AKSHED1; AKA Sprache: Text / Frau (ungefiltert)

H. K. liest einen Textausschnitt von Carola Stern<sup>2</sup> unter Vermeidung längerer Pausen und Unterlassung von „Text-Dramatisierung“. Reproduktion der Lesung von der Tonbandaufnahme. AKA mit 2000 Meßwerten in 300 [s], d.h. in 5 Minuten. Zeitschritt  $\delta t = \delta \tau = 0,15$  [s]; bei 1059 Worten im Mittel 1,7 [Worte/s] gelesen. Die Wortlängen- und Satzlängenanalyse der gesprochenen Worte hat folgendes Ergebnis:

- Wortlänge: Mittelwert 5,87 Buchstaben/Wort ; etwa Poisson-Verteilung mit Maximum bei 3 Buchstaben/Wort; Halbwertsbreite ca. 5 Buchstaben pro Wort.
- Satzlänge: Mittelwert 22,5 Worte/Satz; Maximum 54 Worte/Satz; Minimum 3 Worte/Satz; etwa (breite) Gauß-Verteilung mit Maximum bei 10 Worten/Satz, jedoch etwa gleich große Wahrscheinlichkeit zwischen mit 5 bis 15 Worten/Satz.

Damit ergibt sich eine quasi-periodische Satz-Betonungsmodulation mit der Periodendauer von  $10 \pm 5$  Worten, d.h. von  $[10 \pm 5]/1,7$  [s] =  $\{[10 \pm 5] / [0,15/1,7]\} \cdot \delta t$ . Die Periodendauer beträgt also ca.  $40 \pm 20$  Taktzeiten. Eine Periodendauer dieser Größe von ca.  $20 \cdot \delta \tau$  bis  $60 \cdot \delta \tau$  findet sich erwartungsgemäß auch in der AKF in der Form des Wiederanstiegs ab der Verzögerungszeit  $15 \cdot \delta \tau$ .

In der AKF vorangehend zeigt sich ein quasi-exponentieller Anfangsabfall mit einer geschätzten Zeitkonstanten  $\tau_{1/e} = 4 \cdot \delta \tau = 0,6$  [s] auf einen stationären Wert, der infolge des Wiederanstiegs jedoch nicht sogleich erreicht wird. Dieser exponentielle Abfall korrespondiert genau mit der statistisch schwankenden und unkorreliert auftretenden Wortlänge von im Mittel  $4 \cdot \delta t$ .

### Abb. 6: AKSHED2F; AKA Sprache: Text / Frau ( 6 kHz-Filter )

Gleiche Erläuterungen wie zu Abb. 5 für AKSHED1. Das Tonsignal wurde jedoch innerhalb des Verstärkers durch ein Paß-Filter mit  $f_0 = 6$  kHz und einer 3-dB-Bandbreite von 857 Hz

<sup>2</sup> Carola Stern, „Die Sache, die man Liebe nennt – Das Leben der Fritzi Massary“; rororo-Taschenbuch Nr. 22962, Sonderausgabe 2001; ISBN 3-499-22962-5. Gesprochen wurde der Textausschnitt von Seite 170 (oben) ab mit 1059 Worten, d.h. bis Seite 174 zum Satzende in der Mitte der 8. Zeile von oben.

gefiltert. Damit werden die niederfrequenten quasi-periodischen Modulationsanteile, wie sie durch Satzbetonung entstehen, reduziert gegenüber den 6 kHz-Frequenzanteilen in der Wortaussprache, wie sie insbesondere durch die Zischlaute entstehen.<sup>3</sup>

Die AKF zeigt dementsprechend klarer als in Abb. 5 für AKSHED1 den (gleichen!), auf die Wortlängen-Statistik zurückzuführenden exponentiellen Anfangsabfall mit der Zeitkonstanten  $\tau_{1/e} = 4 \cdot \delta\tau = 0,6$  [s] auf einen quasi-konstanten stationären Endwert.

**Abb. 7: AKMBOL1; Musik: M. Ravel, „Boléro“**

Meßwerte aus der CD-Produktion von André Cluytens mit dem Orchestre de la Societé des Concerts du Conservatoire Paris, © 1990 EMI Electrola GmbH, Köln; CDZ 2522252 / D 537; Gesamtlaufrzeit des Stückes: 15:30; AKA für das Teilstück -7:08 bis -2:10 = 4:58  $\approx$  5 [Min] als „Auswertezeit“; 2000 Meßwerte, d. h. Zeitabstand der Meßwertaufnahme: 0,15 [s]; Handstabilisierung der mittleren Lautstärke; Taktzeit der Aufführung, gestoppt, praktisch konstant in der „Auswertezeit“ mit 0,92 [s/Takt]; Zeitschritt der Verzögerungszeit  $\delta\tau = 0,15$  [s] Periodendauer der AKF: 0,92 [s]; Anfangs-Relaxationszeit der AKF ( ab  $\tau=0$  ):  $\tau_{1/e} \approx 10 \cdot \delta\tau = 1,5$  [s]; Häufigkeitsverteilung der Meßwertgröße: Quasi-Gauß-förmig mit Maximum bei 50 % der insgesamt gegebenen Variation zwischen 0 und 100 %

**Abb. 8: AKSPME5B; periodische Wortwiederholung**

Das Wort „Panamakanal“ wurde von F. R. K. in gleichförmiger Sprechweise quasi-periodisch wiederholt. 1000 Meßwerte mit der Taktzeit  $\delta t = 0,15$  [s] wurden aufgenommen. Periodendauer für ein Wort: im Mittel  $18 \cdot \delta t$ ; Periodendauer des oszillatorischen Anteils der AKF: im Mittel  $16,9 \cdot \delta t$  mit  $\delta\tau = \delta t$

**Abb 9: AKSAT1L; AKA Musik: E. Satie, „Danses gothique“**

Analysiert wurde das Klavierstück von E. Satie: „Danses gotique“, ein frühes Klavierwerk, aus der Langspielplatte Nr. 6.42198 AW, produziert von Harlekijn Holland; Hersteller: ELDEC >>Telefunken-DECCA<< Schallplatten GmbH, 1975. Am Klavier: Reinbert de Leeuw; Gesamtspielzeit: 3'57". Aufnahme von 1500 Meßwerten in 3'25" ; 1 Zeitschritt  $\delta t = 0,138$  [s]. Infolge der rhythmischen Betonung mit einer quasi-konstanten Periodendauer von ca. 4 [s] =  $29 \cdot \delta t$  tritt in der AKF, nach dem anfänglichen exponentiellen Abfall, ebenfalls eine quasi-periodischer Modulationsanteil auf, der sich dem bereits nahezu konstanten stationären Endwertes der AKF überlagert. Dabei ist die Periodendauer in der AKF in Übereinstimmung mit der Betonungsmodulation mit einem Wert von etwa  $29 \cdot \delta\tau$  wegen  $\delta\tau = \delta t$  gleich groß, und entspricht damit der theoretischen Erwartung für die AKF. – Um diese Modulation der AKF länger und damit eindrucksvoller verfolgen bzw. quantitativ ermitteln zu können, ist in der vorliegenden AKA die maximale Verzögerungszeit auf  $300 \cdot \delta\tau$  ausgedehnt worden.

**Abb. 10: AKMFF61A; AKA Musik: ABBA „SOS“ (MW=0; 6 kHz-Filter)**

Analysiert wurde der Schlager „S.O.S.“ der Gruppe ABBA, archiviert auf der Langspiel-Schallplatte: „20 TOP HITS - VOCAL VERSION“, Jahreszeiten-Verlag GmbH (Herausg.), Hamburg, 11/1975; Produktion: „AUDI-TON“; Signal-Selektivfilterung: Mittenfrequenz  $f_0 = 6$  kHz; 3-dB-Bandbreite = 857 Hz; 2170 Meßwerte; Zeitschritt  $\delta t = 0,145$  [s]; Mittelwert der Meßwerte = 0 durch nachträgliche Subtraktion des primären Mittelwertes. Dadurch ist der stationäre Endwert der AKF gleich Null. Normierung der AKF auf den Maximalwert bei  $\tau = 0$ , d.h. auf die Varianz der mittelwertfreien Meßwerte, dadurch  $AKF(\tau = 0) = 1$ . Die Zeitkonstante des primären exponentiellen Abfalls der AKF beträgt etwa  $\tau = 23 \cdot \delta\tau$  bzw. wegen  $\delta\tau = \delta t$  ist  $\tau_{1/e} = 3,4$  [s]. Dieser relativ langsame primäre Abfall wird jedoch alsbald „aufgefangen“ durch einen AKF-Anteil mit Peak-artiger Struktur mit Maximum nach ca.  $30 \cdot \delta\tau$ , d.h. bei etwa

<sup>3</sup> Die Grundtonfrequenzen der menschlichen Stimme erstrecken sich etwa von 45 Hz (= tiefster Ton eines Bassisten) bis ca. 2000 Hz (= höchster Ton einer Sopranistin).



$\tau = 4,5$  [s]. – Diese Art von „Doppelstruktur“ der AKF dürfte einerseits der Existenz eines künstlich, d. h. elektronisch verlängerten Nachhalls und andererseits der rhythmischen Struktur der Musikdarbietung entsprechen.

**Abb. 11: MWSPME1; MW=f(t) Sprache: DLF-Nachr. (Frau)**

Graphische Darstellung der Meßwerte in ihrer zeitlichen Abfolge zu der AKA in Abb. 3. Aufgetragen ist die registrierte Meßwertgröße als Funktion der Meßwert-Nummer der 1000 Meßwerte, deren zeitlicher Abstand  $\delta t = 0,145$  [s] beträgt. Bei der Meßwert-Nr. 330 und 710 erkennt man je eine kleinere Sprechpause infolge „Themenwechsel“. Ansonsten ist der Sprachverlauf relativ „regelmäßig“ und ohne größere „Dramatisierung“, wie es der üblichen Nachrichtenvermittlung im „Deutschlandfunk“ (DLF) entspricht. – Die Wortstruktur ist in der Meßwertabfolge deutlich zu erkennen. – Die zugehörige Häufigkeitsverteilung der Meßwertgröße ist in Abb. 12 dargestellt.

**Abb. 12: HVSPME1; Häufgkt.-Vert. Sprache: DLF-Nachr. (Frau)**

Dargestellt ist die Häufigkeitsverteilung der Meßwertgröße in der registrierten Meßwert-Gesamtheit zu Abb. 11 bzw. zu der entsprechenden AKA in Abb. 3. Sieht man von dem relativ geringen Anteil kleiner Meßwertgrößen ab, so kann man bei der gegebenen Häufigkeitsverteilung von einer relativ schmalen Quasi-Gauß-Verteilung sprechen. Diese Eigenschaft des Signals hat aber keinen primären Einfluß auf die darin enthaltene Korrelation, d.h. auf die AKF des Signals.

**Abb. 13: KOA4M4; eff. Meßwerte=f(X bzw. t); 21.6.01; o. Filter Musik „Bei dir ... schön“; dt=0,18 [s]; stabil. Mittelwert=10**

Zeitliche Abfolge der effektiven Meßwerte bei der „kontinuierlichen Autokorrelationsanalyse“ (vergl. Abschnitt III). Der jeweilige Meßwert ist multiplikativ aus den primären Meßwerten so gebildet, daß bei unveränderter relativer Struktur die Gesamtheit der im Meßwertspeicher abgelegten Meßwerte den Mittelwert der Größe 10 besitzt. Erst nach einem Zeitschritt  $\delta t \equiv \delta t$  mit der Aufnahme eines neuen primären Meßwertes in das Register wiederholt sich der Zyklus der „Mittelwertstabilisierung“ und die Analyse der AKF-Differenz (siehe Abb. 14). Das zugrundeliegende Musikstück ist der konzertant gespielte Schlager „Bei dir war es immer so schön“, es spielt Adolf Steimel und das Organum-Tanzorchester; entnommen der CD „Die besten Tanzorchester von damals“ aus der Serie „Zeitlos“, Nr. 724383812622 der Electrola GmbH, Holland, 1996, LC 0193; Meßwerte ohne Filter aufgenommen,  $\delta t = 0,18$  [s]

**Abb. 14: KOA4R4; Y(X bzw. t) für Verz. Zeit = 7\*dt; 21.6.01; o. Filter Musik „Bei dir ... schön“; Reg.Länge 420; dt = 0,18 [s]**

Zeitabhängigkeit der „relativen normierten Differenz in den Autokorrelationskoeffizienten“ =  $Y(X \text{ bzw. } t)$  für  $\tau_1 = 0$  und  $\tau_2 = 7 * \delta t$  bei der kontinuierlichen Autokorrelationsanalyse nach Abschnitt III; Ordinateneinheit = 1%;  $\delta \tau = dt \equiv \delta t = 0,18$  [s]. Analyse für das in der Legende zu Abb. 13 beschriebene Musikbeispiel. Aufnahme von 2000 effektiven Meßwerten in Zeitschritten von  $\delta t = 0,18$  [s]. Der Mittelwert der Ordinatenwerte beträgt 5,3 %. Entsprechend der Länge des Meßwertregisters von 420 beträgt die Integrationszeit für den Y-Wert  $(420 - 7) * \delta t = 74,34$  [s].

**Abb. 15: KOA4M6; eff. Meßwerte=f(X bzw. t); 22.6.01; o. Filter Sprache „Findelkind“; dt=0,18 [s]; stabil. Mittelwert=10**

Zeitliche Abfolge der effektiven Meßwerte bei der „kontinuierlichen Autokorrelationsanalyse“ (vergl. Abschnitt III). Der jeweilige Meßwert ist multiplikativ aus den primären Meßwerten so gebildet, daß bei unveränderter relativer Struktur die Gesamtheit der im Meßwertspeicher abgelegten Meßwerte den Mittelwert der Größe 10 besitzt. Erst nach einem Zeitschritt  $dt = 0,18$  [s]  $\equiv \delta t$  mit der Aufnahme eines neuen primären Meßwertes in das Register wiederholen sich der Zyklus der „Mittelwertstabilisierung“, die beiden AKAs und die Bestimmung der AKF-Differenz (siehe Abb. 16) Das zugrundeliegende Sprechmuster ist der von F.R.K. vorgelesene



Text: Seite 66 (oben) bis 70 ( ... Pflichterfüllung.) aus: George Sand, „Francois das Findelkind“; manesse im dtv Nr. 24037 Deutscher Taschenbuch Verlag, Manesse Verlag, München, Aug. 1994, ISBN 3-423-24037-7; Aufnahme ohne Spektralfilter am 22.6.01 mit UHER Mikrophon; Schalter auf „Musik“ und Uher-Tonbandgerät 4400 Report Stereo IC; 13-cm-Band „AKA(I)“; Grün / mono: Spur 4; 9,5 cm/s Bandgeschwindigkeit. Analysiert wird ein Ausschnitt aus der Gesamtlaufzeit von ca. 10 Minuten;

**Abb. 16: KOA4R6; Y(x bzw. t) für Verz. Zeit = 7\*dt; 22.6.01; o. Filter; Sprache; „Findelkind“; Reg.Länge 420; dt = 0,18 [s]**

Zeitabhängigkeit der „relativen normierten Differenz in den Autokorrelationskoeffizienten“ =  $Y(X \text{ bzw. } t)$  für  $\tau_1 = 0$  und  $\tau_2 = 7 * \delta\tau$  bei der kontinuierlichen Autokorrelationsanalyse nach Abschnitt III; Ordinateneinheit = 1%;  $\delta\tau = dt \equiv \delta t = 0,18$  [s]. Analyse für das in der Legende zu Abb. 15 beschriebene Sprachbeispiel. Aufnahme von 2000 effektiven Meßwerten in Zeitschritten von  $\delta t = 0,18$  [s]. Der Mittelwert der Ordinatenwerte beträgt 10,8 % Entsprechend der Länge des Meßwertregisters von 420 beträgt die Integrationszeit für den Y-Wert  $(420 - 7) * \delta t = 74,34$  [s].

## Literaturverzeichnis<sup>4</sup>

### A ] Artikel

- [A1] MARIE-LUISE EXNER: Untersuchung unperiodischer Zeitvorgänge mit der Autokorrelations- und der Fourieranalyse, *Acustica*, vol. 4 (1954), p. 365-379
- [A2] K. N. STEVENS: Autocorrelation analysis of speech sounds, *J. Acoust. Soc. Amer.*, vol. 22 (1950), p. 769 # Zitat aus [A1] #
- [A3] Y.W. LEE & J. B. WIESNER: Correlation functions and communication application, *Electronics*, vol. 23 (1950, June), p. 86 # Zitat aus [A1] #
- [A4] S. H. CHANG, G. E. PIHL & M. W. ESSIGMANN: Representations of speech sounds and some of their statistical properties, *Proc. Instn. Radio Engrs.*, vol. 39 (1951), p. 147 # Zitat aus [A1] #
- [A5] W. MEYER-EPPLER: Übersicht über die Verfahren zur Charakterisierung aleatorischer Schallvorgänge und deren Anwendbarkeit auf Geräuschaute, *Z. Phonetik u. allgem. Sprachwissenschaft*, vol. 6 (1952), p. 269 # Zitat aus [A1] #
- [A6] FRANZ RUDOLF KEßLER: Modellierung von Photonenströmen zur vergleichenden Autokorrelationsanalyse, *Abh. der Braunschweigischen Wiss. Gesellschaft*, Band 50 (2001), p. 21-95

### B ] Bücher

- [B1] FREDERIK G. MEE: *Sound*, Verlag William Heinemann, 3. Auflage 1958
- [B2] JOHN E. K. FOREMAN: *Sound analysis and noise control*, Verlag van Nostrand, 1990

---

<sup>4</sup> Zusätze zwischen # ... # enthalten eine Kommentierung durch F. R. Keßler.

- [B3] L. L. BERANEK: Acoustic measurements, Verlag J. Wiley & Sons, New York, 1949  
# Zitat aus [A1] #
- [B4] DIANA V. SHOPOVA & DIMO I. UZUNOV: Correlation, Coherence, and Order, Kluwer Academic / Plenum Publishers, Dordrecht / NL, 1999; ISBN 0-306-46118-8
- [B5] JOHN R. PIERCE: The Science of Musical Sound, Revised Edition, W.H. Freeman and Company Limited, Oxford / England, 1992; ISBN 0-7167-6005-3  
# Die detaillierten Ergebnisse der subtilen Analysen an akustischen Signalen basieren wesentlich auf den Forschungsarbeiten des Autors als „Executive Director Research in the Communication Sciences Division of Bell Telephon Laboratories / USA“, dort auch in Kooperation u. a. mit Manfred Schroeder ( später Direktor des III. Phys. Inst., Uni. Göttingen), und als „Emeritus at Stanford Universitie’s Center for Computer Research in Music and Acoustics / USA“. – Autokorrelationsanalysen an akustischen Signalen werden in dem vorliegenden Buch nicht erwähnt. #
- [B6] J.W.S. RAYLEIGH: The Theory of Sound, 2. Aufl., Bd. 1 - 2; Dover Publ. New York, 1945
- [B7] MICHAEL MÖSER: Analyse und Synthese akustischer Spektren, Springer Verlag, Berlin, 1988; ISBN 3-540-18947-5; zugleich : Teilausgabe der Habil.-Schrift Techn. Universität Berlin

### **C ] Ergänzungen**

- [C1] Z. A. LOMNICKI & S. K. ZAREMBKA: On the estimation of autocorrelation in time series, Ann. Math. Stat., vol. 28 (1967), p. 140-158; ISSN 0003-4851
- [C2] WALTER KRÄMER & RALF RUNDE: Die Autokorrelation von Aktienkursen, Fachbereich Statistik der Universität Dortmund (Herausgeber) Schriftenreihe „Forschungsbericht 90,10“, 1990 ; 25 Seiten
- [C3] ROBERT JOURDAIN: Das wohltemperierte Gehirn - Wie Musik im Kopf entsteht und wirkt, Spektrum Akademischer Verlag, Heidelberg, 2001; ISBN 3-8274-1122-X  
# Der Autor diskutiert zwar die Rolle der Phrasierung als „die vielleicht engste Parallele zwischen Musik und Sprache“ [ auf Seite 336 ], geht auch auf die stil-spezifische Zeitdynamik der Tonfolge und die Bedeutung der „Zwischenspeicherung“ von Ton- bzw. Sprachpassagen im Gehirn für das Verstehen der Inhalte ein, physikalisch-mathematische Begriffe mit Bezug auf Korrelation bzw. Autokorrelation und deren Messung werden jedoch nicht behandelt, d.h. kommen nicht vor. - Das Buch enthält ein umfangreiches Literaturverzeichnis ! #